Particle Density Estimation with Grid-Projected and Boundary-Corrected Adaptive Kernels*

Guillem Sole-Mari[1,2], Diogo Bolster[3], Daniel Fernàndez-Garcia[1,2], Xavier Sanchez-Vila[1,2]

Abstract

The reconstruction of smooth density fields from scattered data points is a procedure that has multiple applications in a variety of disciplines, including Lagrangian (particle-based) models of solute transport in fluids. In random walk particle tracking (RWPT) simulations, particle density is directly linked to solute concentrations, which is normally the main variable of interest, not just for visualization and post-processing of the results, but also for the computation of non-linear processes, such as chemical reactions. Previous works have shown the superiority of kernel density estimation (KDE) over other methods such as binning, in terms of its ability to accurately estimate the "true" particle density relying on a limited amount of information. Here, we develop a grid-projected KDE methodology to determine particle densities by applying kernel smoothing on a pilot binning; this may be seen as a "hybrid" approach between binning and KDE. The kernel bandwidth is optimized locally. Through simple implementation examples, we elucidate several appealing aspects of the proposed approach,

*This work was partially supported by the Spanish Ministry of Economy and Competitiveness through project WE-NEED, PCIN-2015-248.

Email addresses: guillem.sole.mari@upc.edu (Guillem Sole-Mari), dbolster@nd.edu (Diogo Bolster), daniel.fernandez.g@upc.edu (Daniel Fernàndez-Garcia), xavier.sanchez-vila@upc.edu (Xavier Sanchez-Vila)

[1]Department of Civil and Environmental Engineering (DECA), Universitat Politècnica de Catalunya, Barcelona, Spain

[2]Hydrogeology Group (GHS), UPC-CSIC, Barcelona, Spain

[3]Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, South Bend, IN, USA

including its computational efficiency and the possibility to account for typical boundary conditions, which would otherwise be cumbersome in conventional KDE.

Keywords: Particle Density, Adaptive Kernels, Random Walk Particle Tracking, Concentration Estimation, Reactive Transport

1. Introduction

Random Walk Particle Tracking (RWPT) methods are a family of methods commonly used in the hydrologic sciences to simulate transport. They are appealing as they can accurately emulate many different physical processes that occur in natural media such as diffusion, hydrodynamic dispersion, mass transfer across multiple porosity systems and linear sorption [1, 2]. They are also conducive to simulating anomalous non-Fickian transport that arises due to medium heterogeneities below the scale of resolution [e.g., 3]. With RWPTs, the solute mass is discretized into a large number of discrete particles that move across the medium following deterministic and probabilistic rules, which account for the processes of advection, dispersion, matrix diffusion, etc. Lagrangian methods for simulating scalar transport, among which RWPTs are some of the most common, have been shown to be particularly useful when modeling transport in advection-dominated systems, where Eulerian methods can suffer from numerical dispersion and instabilities [1, 4].

However, the main shortcoming of RWPT methods is that, without modification, they may result in very noisy concentration fields due to subsampling effects associated with the finite number of particles in the system. This can be particularly troublesome when simulating solute transport in systems where processes are governed by nonlinearities or tight coupling and interactions between different solute concentrations, of which nonlinear chemical reactions are a prime example. Linear processes such as simple degradation, slow sorption or

chain reactions can efficiently be incorporated to RWPT algorithms by means of additional probabilistic rules with little additional computational cost. On the other hand, nonlinear reactions involve interactions between neighboring particles, which adds complexity to the problemsince particles need to know both their location and the other particles' locations, which can result in an $\mathcal{O}(N^2)$ numerical cost in naive implementations. Even with more optimized approaches that use better search algorithms [e.g., 5], the additional numerical cost can become significant for high particle numbers.

A problem that clearly highlights these issues and has received considerable recent attention in the literature is the simulation of bimolecular reactions of the type $\mathrm{A}+\mathrm{B} \rightarrow \mathrm{C}$ via RWPT [6, 7, 8, 9]. In many such cases it has been shown that the noise associated with the particles can fundamentally change the large scale behavior of the system; in some instances, it may reflect a true noise in the system [10, 11], but in others it may be a numerical artifact that leads to incorrect predictions [12]. Most attempts to simulate other, more complex reactions from a Lagrangian perspective have chosen to attribute concentrations to particles instead of fixed masses [13, 14], and to represent non-advective processes by means of mass transfer, thus allowing arbitrarily complex reactive processes to be simulated on-particle. This approach, however, hinders some of the intrinsic advantages of RWPT. For instance, species-dependent transport properties cannot be readily implemented, at least not in a quick and straightforward manner. Moreover, with these approaches, particles need to be already present at any given location for incoming non-advective solute transport to occur. Hence, since solute mass tends to occupy new fluid as the simulation advances, empty particles need to be included in all those areas where it is anticipated that the solute may reach by dispersion [15].

A first attempt at simulating arbitrarily complex kinetic chemical reactions

with traditional RWPT methods was recently presented by [16]. The method smooths concentration fields using an optimal kernel density estimator. In the paper, the authors derive an expression for the probability of reaction of a particle for any kinetic rate expression, using the optimal kernel density estimator as the particle support volume. This was extended in [17], where using a locally adaptive optimal kernel to determine the solute concentrations and ultimately the particle reaction probabilities provided better results than simple binning, where concentrations are computed by defining fixed representative volumes or bins over which the contained particle mass is assumed to be uniformly distributed. Despite potential shortcomings, the definition of a spatial discretization as a set of fixed bins is very convenient; it allows establishing links between the Lagrangian representation of the solute mass, and other properties that may be defined in space, thus readily enabling nonlinear and coupled interactions. Moreover, particle counting in a regular mesh can be very efficient from a computational viewpoint, typically much more so than a kernel density approach. Thus, one has to weigh the disadvantages and benefits when deciding which approach to take. Ideally, one would have a method designed to obtain the best of all worlds.

Another challenge that often arises in the use of RWPT is the application of nontrivial boundary conditions, best defined in an Eulerian framework. Several methods have been proposed in the recent literature to incorporate different kinds of boundary conditions to RWPT, allowing one to simulate impermeable, Dirichlet [18], or Robin [19, 20, 21] boundary conditions. However, current kernel methods for the reconstruction of the concentrations [e.g., 22] do not address the subject of boundary conditions. Therefore, there is in general no guarantee that the unmodified kernel-based reconstruction will comply with the boundary condition. A trivial example is the case of impermeable boundaries, where,

without correction of the kernel, one may wrongly find nonzero concentration values on the outside of the model domain.

In this paper we propose a robust methodology to obtain solute concentrations from particle positions in RWPT simulations. These concentration fields can then be used to visualize and postprocess the results, and perhaps more importantly, to incorporate nonlinear fate and transport processes. Although we focus on advection-dispersion in relation to a porous medium, this methodology is broadly applicable to other similar transport processes and systems. In fact, any particle-based approach that relies on some form of particle density estimation could benefit from it. Our proposed approach can be seen as a hybrid technique, where the nonreactive part of the transport (advection-dispersion) is simulated following classical principles of RWPT, and the reactive part is decoupled from the former and assisted by a grid on which the concentrations are estimated following an improved form of the local optimal kernel density estimation technique introduced in [17]. This new on-grid kernel-based method combines the practical efficiency of binning techniques with the accuracy gains of kernel methods, while also allowing us to impose Neumann, Dirichlet and Robin boundary conditions to the density estimation.

The paper is structured as follows. First, in §2, we describe the methodology by which we can adapt the kernel density estimation procedure described in [17] to the case of a gridded domain. Then, in §3, we explain how to correct the resulting concentration estimations to account for a variety of boundary conditions typically considered in transport models. Then, in §4 we perform several computational experiments to test and also illustrate the proposed methodology, and in §5 we present the summary and conclusions of the study.

2. On-Grid Concentration Estimation From Particle Positions With Local Optimal Kernels

Let us consider a numerical particle cloud, made up of N particles of identical mass m, that represents the spatial distribution of the total mass of a given chemical compound (solute), over a d-dimensional continuum. Particle positions are given by $\{\mathbf{X}_1, \ldots, \mathbf{X}_N\}$. The concentration of the compound at position $\mathbf{x} \equiv [x_1, \ldots, x_d]^{\mathrm{T}}$ is then expressed as

$$c(\mathbf{x}) = \frac{m\rho(\mathbf{x})}{\phi(\mathbf{x})}, \tag{1}$$

where $\rho(\mathbf{x})$ is the particle density (particles per unit volume of medium), and $\phi(\mathbf{x})$ is the volumetric content of the fluid per unit volume of medium. The particle density $\rho(\mathbf{x})$ needs to be estimated from particle positions, and one method to do so is via kernel density estimation (KDE) such that

$$\rho(\mathbf{x}) \coloneqq \sum_{p=1}^{N} W(\mathbf{x} - \mathbf{X}_p; \mathbf{h}_p), \tag{2}$$

where W is the kernel, chosen here as a "product" multi-Gaussian, defined at distance $\mathbf{r} \equiv [r_1, \ldots, r_d]^{\mathrm{T}}$ from its origin as:

$$W(\mathbf{r}; \mathbf{h}) \coloneqq \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}h_i} \exp\left(-\frac{r_i^2}{2h_i^2}\right), \tag{3}$$

with $\mathbf{h} \equiv [h_1, \ldots, h_d]^{\mathrm{T}}$ being the vector of directional kernel bandwidths. Note that in (2), every particle p can have a different bandwidth $\mathbf{h}_p$. Expression (2) has been used in previous works [22, 23, 12, 16, 17] to link particle positions to solute concentrations. Recently, we [17] proposed a technique to determine the optimal bandwidth $\mathbf{h}_p$ based on the minimization of the root mean squared error (RMSE) on a local environment. In the mentioned paper, the choice of

kernel function given by (3) produced similarly accurate results when compared to other alternatives. Here, we adapt the density estimation and the bandwidth optimization methods to apply them within a regular grid.

2.1. Concentration in a Bin

Let us discretize our entire domain of interest into ν regular bins of size $\boldsymbol{\lambda} \equiv [\lambda_1, \ldots, \lambda_d]^{\mathrm{T}}$, labeled as $u = 1, \ldots, \nu$. Let us also group particle positions into the centers of the containing bins; i.e., if particle p falls into bin u, then $\mathbf{X}_p \approx \mathbf{x}_u$, where $\mathbf{x}_u$ is the position of the center of bin u. Then we can define a discrete (mean) value of the particle density ρ in the uth bin as

$$\rho_u \coloneqq \frac{1}{\Lambda} \int_{\mathbf{x}_u - \boldsymbol{\lambda}/2}^{\mathbf{x}_u + \boldsymbol{\lambda}/2} \rho\left(\mathbf{x}\right) \mathrm{d}\mathbf{x} \approx \frac{1}{\Lambda} \sum_{\omega=1}^{\nu} \mu_\omega \overline{W}\left(\mathbf{x}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right), \tag{4}$$

where $\Lambda = \prod_{i=1}^{d} \lambda_i$ is the measure of the bin, μ_ω is the particle count (number of particles) in bin ω, $\mathbf{h}_\omega$ is its associated kernel bandwidth, and

$$\overline{W}\left(\mathbf{r}; \mathbf{h}, \boldsymbol{\lambda}\right) \coloneqq \int_{\mathbf{r} - \boldsymbol{\lambda}/2}^{\mathbf{r} + \boldsymbol{\lambda}/2} W\left(\mathbf{r}'; \mathbf{h}\right) \mathrm{d}\mathbf{r}'. \tag{5}$$

The kernel $\overline{W}$ is a projected form of W. By combining (3) and (5) we obtain the closed form:

$$\overline{W}\left(\mathbf{r}; \mathbf{h}, \boldsymbol{\lambda}\right) = \frac{1}{2^d} \prod_{i=1}^{d} \left[\operatorname{erf}\left(\frac{r_i + \lambda_i/2}{\sqrt{2} h_i} \right) - \operatorname{erf}\left(\frac{r_i - \lambda_i/2}{\sqrt{2} h_i} \right) \right]. \tag{6}$$

Note from (4) that, since all $\mathbf{x}_u$, $\mathbf{x}_\omega$ belong to a regular grid of size $\boldsymbol{\lambda}$, then $\mathbf{r}$ in (6) can be written as

$$\mathbf{r} = \boldsymbol{\lambda} \odot \mathbf{z}, \tag{7}$$

where $\mathbf{z} = [z_1, \ldots, z_d]^{\mathrm{T}}$ is a vector of integers ($z_i \in \mathbb{Z}$, $\forall i = 1, \ldots, d$), and $\odot$ is the Hadamard (element-wise) product. That is, $\mathbf{z}$ is a cell index, with $\mathbf{z} = \mathbf{0}$

corresponding to the cell where the kernel is centered. Then expression (6) can be rewritten as

$$\overline{W}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{h},\boldsymbol{\lambda}\right)=\frac{1}{2^d}\prod_{i=1}^{d}\left[\operatorname{erf}\left(\lambda_i\frac{z_i+1/2}{\sqrt{2}h_i}\right)-\operatorname{erf}\left(\lambda_i\frac{z_i-1/2}{\sqrt{2}h_i}\right)\right]. \tag{8}$$

By examining equation (8), we see that for any given ratio $\mathbf{h}\oslash\boldsymbol{\lambda}$ (where $\oslash$ is the Hadamard division) the set of required evaluations of $\overline{W}$ in (4) (setting a cutoff distance) can be fully defined as a d-dimensional matrix (see leftmost illustrations in Figure 1). Then, by limiting the possible values that $\mathbf{h}$ can adopt to a discrete set, it is possible to greatly speed up evaluation of (4) by storing these matrices after their first computation, hence never having to re-evaluate (6) for similar values of $\mathbf{h}_\omega$. More details on the properties, generation, cut-off correction, storage and use of matrix kernels can be found in Appendix A.

Now, the kernel-based evaluation of particle densities on bins consists of two steps. First, a simple binning is performed to obtain the particle count μ_ω in every bin, which can be seen as an initial, perhaps noisy, "pilot" density estimation. Then, smoothing (4) is performed using the matrix kernels corresponding to bandwidths $\mathbf{h}_\omega$ to obtain densities ρ_u.

Finally, the concentration in bin u can be calculated as

$$c_u\coloneqq\frac{1}{\Lambda}\int_{\mathbf{x}_u-\boldsymbol{\lambda}/2}^{\mathbf{x}_u+\boldsymbol{\lambda}/2}c\left(\mathbf{x}\right)\mathrm{d}\mathbf{x}\approx\frac{m\rho_u}{\phi_u}, \tag{9}$$

where $\phi_u\coloneqq\phi\left(\mathbf{x}_u\right)$. Details on the local selection of parameter $\mathbf{h}_\omega$ to use in (4) are given in the following section.

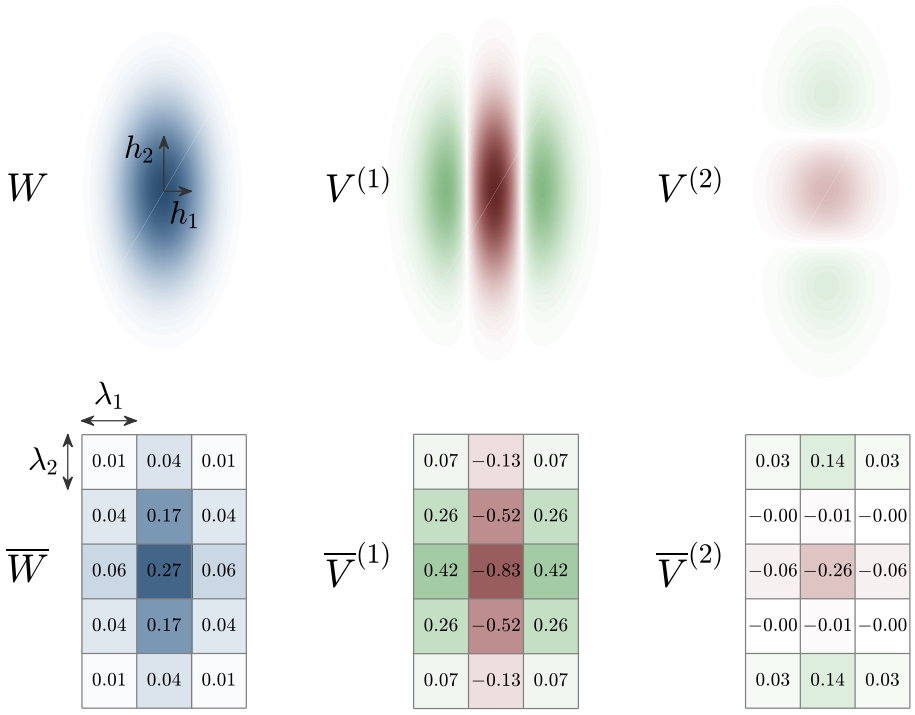

Figure 1: Graphical example of 2D kernel functions (top) and their discrete equivalents obtained by projection on a Eulerian grid (bottom). Color intensity is proportional to the kernel density, and numbers correspond to the integral over a cell ($\Lambda\rho_u$).

2.2. The Optimal Kernel Bandwidth $\mathbf{h}$

Let us decompose the bandwidth $\mathbf{h}$ by defining the bandwidth scale $\widehat{h}$ and the vector of "shape" parameters $\mathbf{s} \equiv [s_1, \ldots, s_d]^{\mathrm{T}}$:

$$\widehat{h} \coloneqq \left(\prod_{i=1}^{d} h_i \right)^{\frac{1}{d}}, \qquad s_i \coloneqq h_i / \widehat{h}, \tag{10}$$

i.e., $\mathbf{h} \equiv \widehat{h}\mathbf{s}$, with $\prod_{i=1}^{d} s_i = 1$. Then, following [17], the local optimal bandwidth scale is

$$\widehat{h}_u = \left[\frac{d\ n_u}{(4\pi)^{\frac{d}{2}}\ T_u} \right]^{\frac{1}{d+4}}, \tag{11}$$

where

$$n_u \coloneqq \int \rho\left(\mathbf{x}\right) W\left(\mathbf{x}-\mathbf{x}_u;\boldsymbol{\sigma}_u\right)\mathrm{d}\mathbf{x} \approx \sum_{\omega=1}^{\nu} \rho_\omega \overline{W}\left(\mathbf{x}_\omega-\mathbf{x}_u;\boldsymbol{\sigma}_u,\boldsymbol{\lambda}\right), \tag{12}$$

with $\boldsymbol{\sigma}_u$ being the bandwidth of a local integration kernel (see §2.3), and

$$T_u = \begin{cases} \Psi_{11,u}, & \text{for } d=1, \\ 2\left(\Psi_{11,u}\Psi_{22,u}\right)^{\frac{1}{2}} + 2\Psi_{12,u}, & \text{for } d=2, \\ 3\left(\Psi_{11,u}\Psi_{22,u}\Psi_{33,u}\right)^{\frac{1}{3}} + \sum_{i\neq j\neq k}\Psi_{ij,u}\left(\dfrac{\Psi_{kk,u}^2}{\Psi_{ii,u}\Psi_{jj,u}}\right)^{\frac{1}{6}}, & \text{for } d=3. \end{cases} \tag{13}$$

In (13), the functionals $\Psi_{ij,u}$ are defined as

$$\begin{aligned} \Psi_{ij,u} \coloneqq \int \kappa^{(i)}\left(\mathbf{x}\right)\kappa^{(j)}\left(\mathbf{x}\right) W\left(\mathbf{x}-\mathbf{x}_u;\boldsymbol{\sigma}_u\right)\mathrm{d}\mathbf{x} \\ \approx \sum_{\omega=1}^{\nu} \kappa_\omega^{(i)}\kappa_\omega^{(j)}\overline{W}\left(\mathbf{x}_\omega-\mathbf{x}_u;\boldsymbol{\sigma}_u,\boldsymbol{\lambda}\right), \end{aligned} \tag{14}$$

where

$$\kappa^{(i)}\left(\mathbf{x}\right) \coloneqq \frac{\partial^2\rho}{\partial x_i^2}\left(\mathbf{x}\right) = \sum_{p=1}^{N} V^{(i)}\left(\mathbf{x}-\mathbf{X}_p;\mathbf{g}_p^{(i)}\right), \tag{15}$$

with the kernel function $V^{(i)}$ being defined as

$$V^{(i)}\left(\mathbf{r};\mathbf{g}\right) \coloneqq \frac{\partial^2 W}{\partial r_i^2} = \left(\frac{r_i^2}{g_i^4}-\frac{1}{g_i^2}\right) W\left(\mathbf{r};\mathbf{g}\right). \tag{16}$$

In (14), $\kappa_u^{(i)}$ is defined as the mean value of $\kappa^{(i)}$ in bin u:

$$\kappa_u^{(i)} \coloneqq \frac{1}{\Lambda}\int_{\mathbf{x}_u-\boldsymbol{\lambda}/2}^{\mathbf{x}_u+\boldsymbol{\lambda}/2} \kappa^{(i)}\left(\mathbf{x}\right)\mathrm{d}\mathbf{x} \approx \frac{1}{\Lambda}\sum_{\omega=1}^{\nu}\mu_\omega\overline{V}^{(i)}\left(\mathbf{x}_\omega-\mathbf{x}_u;\mathbf{g}_\omega^{(i)},\boldsymbol{\lambda}\right), \tag{17}$$

with

$$\overline{V}^{(i)}\left(\mathbf{r};\mathbf{g},\boldsymbol{\lambda}\right) \coloneqq \int_{\mathbf{r}-\boldsymbol{\lambda}/2}^{\mathbf{r}+\boldsymbol{\lambda}/2} V^{(i)}\left(\mathbf{r}';\mathbf{g}\right)\mathrm{d}\mathbf{r}'. \tag{18}$$

The notation $\mathbf{g}_\omega^{(i)}$ *in lieu* of $\mathbf{h}_\omega$ indicates that this bandwidth can adopt a different value than that of kernel W, and also different values for different directions of derivative. By combining (16) and (18) we obtain the following closed form:

$$\overline{V}^{(i)}(\mathbf{r};\mathbf{g},\boldsymbol{\lambda}) = -\frac{1}{2^{(d-\frac{1}{2})}\sqrt{\pi}g_i^3}\times \left[\left(r_i+\frac{\lambda_i}{2}\right)\exp\left(-\frac{(r_i+\lambda_i/2)^2}{2g_i^2}\right)-\left(r_i-\frac{\lambda_i}{2}\right)\exp\left(-\frac{(r_i-\lambda_i/2)^2}{2g_i^2}\right)\right]\times \prod_{j\neq i}\left[\operatorname{erf}\left(\frac{r_j+\lambda_j/2}{\sqrt{2}g_j}\right)-\operatorname{erf}\left(\frac{r_j-\lambda_j/2}{\sqrt{2}g_j}\right)\right]. \tag{19}$$

Similar to $\overline{W}$, the quantity $\lambda_i^2\overline{V}^{(i)}$ can be stored as a matrix (Figure 1) with values that only depend on the ratio $\mathbf{g}\oslash\boldsymbol{\lambda}$. Additional corrections are performed on kernel $\overline{V}^{(i)}$ to ensure that it keeps the main properties of the original kernel $V^{(i)}$ despite the projection. Details on these corrections, as well as the generation, storage and use of matrix kernels can be found in Appendix A.

The shape vector is also determined by the "roughness" functionals $\Psi_{ij,u}$:

$$s_{i,u} = \left(\frac{\widehat{\Psi}_u}{\Psi_{ii,u}}\right)^{\frac{1}{4}}, \qquad \widehat{\Psi}_u \coloneqq \left(\prod_{j=1}^{d}\Psi_{jj,u}\right)^{\frac{1}{d}}. \tag{20}$$

Detailed information on the theory behind these expressions can be found in [17]. Nevertheless, here we give an intuitive explanation: In (11), n_u is a smooth average of the particle density over a local environment, whereas T_u, termed as "effective roughness", is a measure of the square of the second spatial derivatives of the densities, also averaged over a local environment. Thus, typically, those areas where particle densities are higher, or where they form "peaks" and "valleys" that are more pronounced, will yield smaller kernel bandwidths (lower values of $\widehat{h}_u$). In (20), we see that the elongation of the kernel bandwidth in a

direction i will be inversely correlated to the (region-averaged) squared derivatives of the density in that direction, normalized by its geometric average over all dimensions; that is, the kernel will tend to stretch along the direction of minimum curvature. In sum, the local kernel's size and shape adapts to "mimic" the features of the surrounding particle cloud.

At this juncture, the computation of the optimal bandwidth $\mathbf{h}_u$ to use in (4) requires the input of two additional kernel bandwidths: $\boldsymbol{\sigma}_u$ in (12) and (14) ; and $\mathbf{g}_\omega^{(i)}$ in (17). These are addressed in the following two sections, respectively.

2.3. The Integration Support $\boldsymbol{\sigma}$

Computation of (12) and (14) requires the definition of a Gaussian integration support, represented here by its vector of directional bandwidths $\boldsymbol{\sigma}_u$. In the original development of the local optimization methodology [17], an isotropic, spatially constant support $\sigma_i \equiv \widehat{\sigma}$, $\forall i = 1, \ldots, d$ was proposed, such that $\widehat{\sigma} = 3\widehat{h}^{\mathcal{G}}$, with $\widehat{h}^{\mathcal{G}}$ defined as the global AMISE-optimal kernel bandwidth scale. Not only is this approach completely heuristic in nature, but it also renders the local kernel indirectly dependent on a global feature, compromising local benefits. Here, we overcome this problem by introducing the concept of an equivalent normal particle distribution. We assume that the local optimal kernel scale $\widehat{h}_u$ is also the global optimal kernel bandwidth $\widehat{h}_u^\sigma$ associated with a virtual Gaussian distribution of variance $\widehat{\sigma}_u^2$, composed of a virtual number of particles N_u^σ; the classic expression for the AMISE-optimal kernel bandwidth in this case is [24]:

$$\widehat{h}_u \equiv \widehat{h}_u^\sigma = \left[\frac{4}{(d+2)\, N_u^\sigma}\right]^{\frac{1}{d+4}} \widehat{\sigma}_u. \tag{21}$$

We then impose that this virtual distribution has locally matching values with the actual particle distribution for ρ_u and n_u, the latter being defined as in (12), using $\boldsymbol{\sigma}_u$ as the integration support. Then, it can be shown (see Appendix B)

that the following relation holds,

$$N_u^{\sigma} = \frac{\left(\sqrt{8\pi}\widehat{\sigma}_u\right)^d n_u^2}{\rho_u}. \tag{22}$$

Combining (21) with (22), and after some algebraic manipulation, we obtain:

$$\widehat{\sigma}_u = \left[\frac{(d+2)\,(8\pi)^{\frac{d}{2}}\, n_u^2 \widehat{h}_u^{d+4}}{4\rho_u}\right]^{\frac{1}{4}}. \tag{23}$$

The integration support used in (12) and (14) is then $\boldsymbol{\sigma}_u = \widehat{\sigma}_u \mathbf{1}$, where $\mathbf{1}$ is a $d \times 1$ vector of ones. Note that expression (23) is recursive, in the sense that h_u, n_u and ρ_u are all affected by some previous choice of $\widehat{\sigma}_u$. Nonetheless, it can be implemented iteratively. More details on the iterative implementation are given in §2.5.

2.4. The Curvature Kernel Bandwidth $\mathbf{g}^{(i)}$

Computation of the optimal bandwidth as presented in §2.2 requires a bandwidth for the estimation of the particle density curvatures ($\mathbf{g}_{\omega}^{(i)}$ in (15)). In [17], the Improved Sheather-Jones plug-in method by *Botev et al.* [25] was used to determine this bandwidth. Here, we rely on the equivalent Gaussian particle distribution described in §2.3 to determine a local value for $\mathbf{g}_{\omega}^{(i)}$ recursively based on $\mathbf{h}_{\omega}$.

Above, we have defined both these bandwidths as diagonal. Here we further assume that $\mathbf{g}^{(i)}$ is isotropic,

$$g_j^{(i)} \equiv \widehat{g}^{(i)},\ \forall j = 1, \ldots, d. \tag{24}$$

We also assume that the anisotropy of $\boldsymbol{\sigma}_u$ is identical to that of bandwidth $\mathbf{h}_u$ (represented by vector $\mathbf{s}_u$). Then, in a similar fashion to what was done in §2.3 for $\widehat{h}_u$, the value of bandwidth scale $\widehat{g}_u^{(i)}$ is assumed to match its AMISE-optimal

magnitude within the virtual Gaussian distribution of particles defined in §2.3 through $\boldsymbol{\sigma}_u$ and N_u^σ. This magnitude (see derivation in Appendix C) is

$$\widehat{g}_u^{(i)} \equiv \widehat{g}_u^{(i),\sigma} = \left[\frac{4 + 2^{\frac{d}{2}+4}}{3\,(d+4)\,N_u^\sigma}\right]^{\frac{1}{d+6}} \widehat{\sigma}_u \vartheta_i\,(\mathbf{s}_u)\,, \tag{25}$$

where ϑ_i is a function of the anisotropy of $\mathbf{s}_u$, with a value of 1 in the isotropic case (i.e., $s_i = 1,\ \forall i = 1, \ldots, d$):

$$\vartheta_i\,(\mathbf{s}) = \left[\frac{1}{d+4}\sum_{j=1}^{d}\frac{1+4\delta_{ij}}{s_i^4 s_j^2}\right]^{-\frac{1}{d+6}}, \tag{26}$$

where δ_{ij} is the Kronecker delta. By combining (25) and (21), we can determine the ratios between optimal bandwidth scales:

$$\gamma_u^{(i)} \coloneqq \frac{\widehat{g}_u^{(i)}}{\widehat{h}_u} = \alpha \cdot (N_u^\sigma)^\beta \cdot \vartheta_i\,(\mathbf{s}_u)\,, \tag{27}$$

$$\alpha \coloneqq \left[\frac{1 + 2^{\frac{d+4}{2}}}{3 \cdot 2^{\frac{4}{d+4}}}\right]^{\frac{1}{d+6}} \cdot \frac{(d+2)^{\frac{1}{d+4}}}{(d+4)^{\frac{1}{d+6}}}, \qquad \beta \coloneqq \frac{2}{(d+4)\,(d+6)}, \tag{28}$$

with N_u^σ obtained from (22). It is worth noting that α in (27) is a constant value close to 1 (ranging from 1.04 for $d = 1$ to 1.12 for $d = 3$), and $\beta \ll 1$ (meaning that the ratio $\gamma_u^{(i)}$ is very rigid with respect to N_u^σ). Hence $\widehat{g}_u^{(i)}$ is typically just somewhat larger than $\widehat{h}_u$ for a wide range of values of N_u^σ. For instance, in 2D, assuming isotropy ($\vartheta_i = 1$), $\gamma_u^{(i)} \approx 1.31$ for $N_u^\sigma = 10^2$, and $\gamma_u^{(i)} \approx 1.75$ for $N_u^\sigma = 10^5$. Using definition (27), we can recursively obtain $\widehat{g}_u^{(i)} = \gamma_u^{(i)}\widehat{h}_u$, and then, in equation (17), we use $\mathbf{g}_u = \widehat{g}_u\mathbf{1}$.

2.5. Optimization Algorithm

The approach that we propose to determine the locally optimal bandwidth can be seen as a fixed-point iteration method: A set of local kernel bandwidths

is given as an input, and a new set is obtained as an output, until convergence. Below, for any variable "a", the notation $\{a\}$ indicates the set of all a_u. Before starting the iteration process, the particle counts $\{\mu\}$ are computed, and, if unavailable, an initial pilot $\{\widehat{\sigma}\}$ is defined such that $\widehat{\sigma}_u = 3\widehat{h}_u$. Then, the structure of one iteration can be summarized as follows:

1. Compute $\{\rho\}$ via (4) using the input $\{\mathbf{h}\}$.

2. To obtain $\{\widehat{\sigma}\}$ and $\{n\}$:

   Compute pilot $\{n\}$ via (12) using $\{\rho\}$ and $\{\widehat{\sigma}\}$.

   Compute $\{\widehat{\sigma}\}$ via (23) using $\{\widehat{h}\}$, $\{\rho\}$ and pilot $\{n\}$.

   Compute $\{n\}$ via (12) using $\{\rho\}$ and $\{\widehat{\sigma}\}$.

3. For $i = 1, \ldots, d$:

   Compute $\{\widehat{g}^{(i)}\}$ via (27) using $\{\rho\}$, $\{n\}$, $\{\widehat{\sigma}\}$ and $\{\mathbf{s}\}$.

   Compute $\{\kappa^{(i)}\}$ via (17) using $\{g^{(i)}\}$.

4. For $i = 1, \ldots, d$, for $j = i, \ldots, d$:

   Compute $\{\Psi_{ij}\}$ via (14) using $\{\widehat{\sigma}\}$, $\{\kappa^{(i)}\}$ and $\{\kappa^{(j)}\}$.

5. Compute the new $\{\mathbf{h}\}$ via (11) and (20) using $\{n\}$ and all $\{\Psi_{ij}\}$.

The iteration process may be exited when relative changes in $\{\mathbf{h}\}$ are below some tolerance level. In the context of a full RWPT simulation, the local kernel bandwidths will evolve in time following the deformation of the particle plume. Since changes in the optimal kernel bandwidth typically occur at a much slower pace than solute transport [17], the optimization does not need to be performed at every time step of the transport time discretization, and if performed often enough, a single iteration may suffice. Between optimizations, particles should "carry" the identifier of the bin they were located in at the time of the latest

optimization; then, the bandwidth at a bin is the average of the bandwidths associated to the contained particles.

3. Boundary Condition Corrections

Kernel methods typically suffer from boundary bias problems [e.g., 26]. Near no-flux boundaries, the standard KDE as written in (2) may for instance produce an underestimation bias; as its support can cross the boundary in question, it may unphysically project nonzero solute concentrations on the other side of such boundaries, where no such mass can actually exist. Similar problems would occur for constant concentration or fixed flux boundaries, where without some correction, kernels will project incorrect and unphysical concentrations at and across the boundaries. Such problems also arise with the proposed discrete estimator in (4). However, for the discrete case, the position of the bins with respect to the boundaries is known and does not change over the course of a simulation. This allows us to efficiently introduce corrections to account for the presence and influence of boundaries, that would otherwise be unfeasible (or at least extremely cumbersome) directly from (2).

Here we propose an approach based on the assumption that the Gaussian kernel that represents a particle's support volume behaves like a purely diffusive process, whose interactions with boundaries are well understood. In other words, we treat the kernel associated with a particle as if it were the result of a very fast diffusive process that has resulted in the spatial spreading of the particle's mass, with an initial condition of a Dirac delta located at the particle's physical position. If the particle is far from the boundary, this results in the standard Gaussian kernel (3). Previous works [e.g., 25] acknowledge this link between KDE and diffusion. The correction on the kernel that we propose is independent of the implementation of the boundary condition itself on the

RWPT algorithm [see 18]; that is, the particles undergoing the random walks in the system are unaware of the kernel and so their motion must still account for the presence of a boundary (e.g., reflection on a no-flux boundary). Note also that, in principle, our proposed approach could be extended to any kernel-based Lagrangian method [e.g., 27] in a bounded domain with physical boundary conditions. For most commonly used boundary conditions, we derive the simple reflection principles that can be used to modify the kernel near a regular boundary that is aligned with the principal directions of the kernel. Then, we also propose an approach to extend this procedure to the case of irregularly-shaped boundaries, which present unique challenges.

3.1. Regular Boundaries

All boundary conditions most typically used in transport simulations can be written in terms of a balance of mass fluxes between the inner (+) and the outer (−) side of the boundary; i.e.:

$$\mathbf{n}^{\mathrm{T}}\left[(\phi\mathbf{D}\boldsymbol{\nabla}c)^{+}-(\phi\mathbf{D}\boldsymbol{\nabla}c)^{-}\right]=\mathbf{n}^{\mathrm{T}}\left[(\mathbf{q}c)^{+}-(\mathbf{q}c)^{-}\right], \tag{29}$$

where $\mathbf{n}$ is the unit vector normal to the boundary. Depending on the assumptions made for the outer side of the boundary, this balance can result in a variety of common boundary conditions.

3.1.1. Impermeable/Outlet

If the boundary is assumed to be a no-flux boundary (impermeable), we set $\mathbf{n}^{\mathrm{T}}\mathbf{q}=0$, and $(\phi\mathbf{D}\boldsymbol{\nabla}c)^{-}=0$. In this case, equation (29) becomes the homogeneous Neumann boundary condition, which, for boundary $\mathfrak{N}$, can be written as

$$\mathbf{n}^{\mathrm{T}}\left(\mathbf{x}_{\mathfrak{N}}\right)\boldsymbol{\nabla}c\left(\mathbf{x}_{\mathfrak{N}}\right)=0,\qquad\mathbf{x}_{\mathfrak{N}}\in\mathfrak{N}. \tag{30}$$

Note that (30) would also be obtained at an outlet, by assuming $\mathbf{q}c^- = \mathbf{q}c^+$, and $(\phi \mathbf{D} \boldsymbol{\nabla} c)^- = 0$.

Let us consider a particle p located at $\mathbf{X}_p$. The density kernel associated with this particle is $W(\mathbf{x} - \mathbf{X}_p; \mathbf{h}_p)$, if we neglect the effect of the boundary. As noted above, let us assume that this density distribution can be seen as the Green's function of a virtual, fast diffusive process. Then, if the boundary is regular and aligned with the principal directions of the kernel, the kernel is altered by the proximity to $\mathfrak{N}$ by a reflection principle such that

$$W_{\mathfrak{N}}(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p) \coloneqq W(\mathbf{x} - \mathbf{X}_p; \mathbf{h}_p) + W(\mathbf{x} - \widetilde{\mathbf{X}}_p; \mathbf{h}_p), \tag{31}$$

where $\widetilde{\mathbf{X}}_p$ is the mirror of $\mathbf{X}_p$ through $\mathfrak{N}$:

$$\widetilde{\mathbf{X}}_p \coloneqq \mathbf{X}_p - 2\mathbf{n}\mathbf{n}^{\mathrm{T}}(\mathbf{X}_p - \mathbf{x}_{\mathfrak{N}}). \tag{32}$$

Because of the regularity assumption,

$$|\mathbf{x} - \widetilde{\mathbf{X}}_p| \equiv |\widetilde{\mathbf{x}} - \mathbf{X}_p|, \tag{33}$$

where

$$\widetilde{\mathbf{x}} \coloneqq \mathbf{x} - 2\mathbf{n}\mathbf{n}^{\mathrm{T}}(\mathbf{x} - \mathbf{x}_{\mathfrak{N}}), \tag{34}$$

which allows the corrected density estimation (i.e. replacing W with $W_{\mathfrak{N}}$ in (2)) to be expressed as a function of the uncorrected density estimation:

$$\rho_{\mathfrak{N}}(\mathbf{x}) \coloneqq \sum_{p=1}^{N} W_{\mathfrak{N}}(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p) = \rho(\mathbf{x}) + \rho(\widetilde{\mathbf{x}}). \tag{35}$$

Hence, densities can be computed conventionally by (2), then the corrected density at a point is obtained by adding the density for the mirror symmetry

point. It follows directly, that for the on-grid methodology presented in this work,

$$\rho_{\mathfrak{N},u} = \rho_u + \rho_{\widetilde{u}}, \tag{36}$$

where $\rho_{\widetilde{u}}$ is the uncorrected density (4) computed at the mirror bin $\widetilde{u}$, whose center is located at

$$\widetilde{\mathbf{x}}_u = \mathbf{x}_u - 2\mathbf{n}\mathbf{n}^{\mathrm{T}} \left(\mathbf{x}_u - \mathbf{x}_{\mathfrak{N}}\right). \tag{37}$$

3.1.2. Dirichlet

Assuming in (29) that the dispersive flux, $\phi\mathbf{D}\boldsymbol{\nabla}c$, is the same on the outer and inner sides of the boundary, and that the outer concentration is prescribed, we obtain a Dirichlet boundary condition,

$$c\left(\mathbf{x}_{\mathfrak{D}}\right) = c_{\mathrm{o}}, \tag{38}$$

where c_{o} is the prescribed concentration. In this case, the boundary can be permeable to a diffusive process. Since this is true in both directions, part of a particle's mass may fall outside the domain, and mass may enter the domain as well. Based again on the assumption that the kernel represents a fast diffusive process, we can separate this process into two parts following the superposition principle, first accounting for the initial condition with a homogeneous boundary condition (a perfectly absorbing boundary), and second accounting for the inhomogeneous part:

$$\begin{aligned} \rho_{\mathfrak{D}}\left(\mathbf{x}\right) &= \rho_{\mathfrak{D}\mathrm{H}}\left(\mathbf{x}\right) + \rho_{\mathfrak{D}\mathrm{I}}\left(\mathbf{x}\right) \\ &\equiv \sum_{p=1}^{N} W_{\mathfrak{D}\mathrm{H}}\left(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p\right) + \sum_{p=1}^{N} W_{\mathfrak{D}\mathrm{I}}\left(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p\right). \end{aligned} \tag{39}$$

The solution to the homogenous part, $W_{\mathfrak{D}\mathrm{H}}$, given a regular boundary aligned with the principal directions of the kernel, is [28]:

$$W_{\mathfrak{D}\mathrm{H}}\left(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p\right) = W\left(\mathbf{x} - \mathbf{X}_p; \mathbf{h}_p\right) - W(\mathbf{x} - \widetilde{\mathbf{X}}_p; \mathbf{h}_p). \tag{40}$$

The inhomogeneous part can also be solved under the same assumptions. Constant diffusion from a regular boundary follows the 1D analytical solution in the direction normal to the boundary [29]:

$$\rho_{\mathfrak{D}\mathrm{I}}\left(\mathbf{x}; \mathbf{h}\right) = \rho_{\mathrm{o}}\mathrm{Erfc}\left(\frac{1}{\sqrt{2}}\mathbf{n}\left[\left(\mathbf{x} - \mathbf{x}_{\mathfrak{D}}\right) \oslash \mathbf{h}\right]\right), \tag{41}$$

with

$$\rho_{\mathrm{o}} \coloneqq \frac{\phi\left(\mathbf{x}_{\mathfrak{D}}\right) c_{\mathrm{o}}}{m}, \tag{42}$$

assuming a constant $\phi \approx \phi\left(\mathbf{x}_{\mathfrak{D}}\right)$ near the boundary. Expression (41) is also the Green's function of diffusion for an initial condition of uniform density $2\rho_{\mathrm{o}}$ at the outer side of the boundary [18]. Thus, we note that equation (41) is approximately equivalent to a sum of mirror kernels (as in (39)) provided that

$$W_{\mathfrak{D}\mathrm{I}}\left(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p\right) = \frac{2\rho_{\mathrm{o}}}{\rho_*(\mathbf{X}_p)} W(\mathbf{x} - \widetilde{\mathbf{X}}_p; \mathbf{h}_p), \tag{43}$$

where ρ_* is the unknown true density. We replace it by a pilot estimate at the containing bin ω, i.e.,

$$W_{\mathfrak{D}\mathrm{I}}\left(\mathbf{x}, \mathbf{X}_p; \mathbf{h}_p\right) \approx \frac{2\mu_{\mathrm{o}}}{\mu_\omega} W(\mathbf{x} - \widetilde{\mathbf{X}}_p; \mathbf{h}_p), \qquad \mathbf{X}_p \approx \mathbf{x}_\omega, \tag{44}$$

with $\mu_{\mathrm{o}} \coloneqq \Lambda\rho_{\mathrm{o}}$. Then, substituting (40) and (44) into (39), and integrating as in (4), the on-grid corrected density estimate becomes:

$$\rho_{\mathfrak{D},u} = \rho_u + \rho^{\mathrm{o}}_{\widetilde{u}}, \tag{45}$$

where

$$\rho^{\mathrm{o}}_{\widetilde{u}} \coloneqq \frac{1}{\Lambda}\sum_{\omega=1}^{\nu}\left(2\mu_{\mathrm{o}} - \mu_\omega\right)\overline{W}\left(\mathbf{x}_\omega - \widetilde{\mathbf{x}}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right), \tag{46}$$

where mirror bin $\widetilde{u}$ is defined in analogy with (37). Expression (45) is somewhat similar to (36); however now the simple reflection $\rho_{\widetilde{u}}$ is replaced by a new term, $\rho^{\mathrm{o}}_{\widetilde{u}}$. Examining expression (46) we see that the only modification in the smoothing process concerns the transfer of density from any non-empty bin ω to an external bin $\widetilde{u}$: it must be done as if the bin ω contained a virtual number of particles equal to $2\mu_{\mathrm{o}} - \mu_\omega$, instead of the actual value of μ_ω.

3.1.3. Robin

If we assume in (29) that the outer side of the boundary is a reservoir with a prescribed concentration, i.e., $\left(\phi\mathbf{D}\boldsymbol{\nabla}c\right)^- = 0$ and $c^- = c_{\mathrm{o}}$, what we obtain is a Robin (or third-type) boundary condition, which can be written for boundary $\mathfrak{R}$ as

$$\mathbf{n}^{\mathrm{T}}\left(\mathbf{x}_{\mathfrak{R}}\right)\mathbf{D}\left(\mathbf{x}_{\mathfrak{R}}\right)\boldsymbol{\nabla}c\left(\mathbf{x}_{\mathfrak{R}}\right) = \mathbf{n}^{\mathrm{T}}\left(\mathbf{x}_{\mathfrak{R}}\right)\mathbf{v}\left(\mathbf{x}_{\mathfrak{R}}\right)\left(c\left(\mathbf{x}_{\mathfrak{R}}\right) - c_{\mathrm{o}}\right), \qquad \mathbf{x}_{\mathfrak{R}} \in \mathfrak{R} \tag{47}$$

where $\mathbf{v} = \mathbf{q}/\phi$. Like in the previous cases, we treat the kernel as a very fast "diffusive" process, which interacts with the boundary analogously to the Green's function of the simulated dispersion. Since, for very large $\mathbf{D}$, condition (47) converges to (30), a Robin boundary affects the kernel identically to a

homogeneous Neumann boundary (see §3.1.1):

$$\rho_{\mathfrak{R},u} = \rho_u + \rho_{\widetilde{u}}. \tag{48}$$

This would also apply to similar Robin boundary conditions originating from other assumptions, such as reactive walls [e.g., 30].

3.2. Extension to Irregular Boundaries

The reflection principles given in §3.1 are derived based on two assumptions: (i) the boundaries are regular and (ii) they are aligned with the principal directions of the kernel. The latter can always be fulfilled by using an isotropic kernel (restricting the degrees of freedom of the bandwidth, see [17]). However, the assumption of regular boundaries may just not be fulfilled, and then expressions (36), (45) and (48) might not be valid as they rely on substituting the mirroring of the kernel for that of the evaluation point through (33). Moreover, in that case, the true Green's function of the diffusion process associated to bandwidth $\mathbf{h}_\omega$ is much more complex than the solutions given by (31), (40) and (43). Nevertheless, as an approximation, one can use these solutions to modify the kernel directly, defining the mirror points by reflection through the closest boundary point, and with a weighting parameter to compensate for the irregularity. For impermeable or Robin, we propose the following kernel:

$$\overline{W}_{\mathfrak{R}}\left(\mathbf{x}_\omega, \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right) \coloneqq \overline{W}\left(\mathbf{x}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right) + \eta_\omega \overline{W}\left(\widetilde{\mathbf{x}}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right), \tag{49}$$

where η_ω is a weighting parameter such that the mass of the "reflection" kernel $\overline{W}\left(\widetilde{\mathbf{x}}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right)$ entering the domain matches the mass of the "regular" kernel $\overline{W}\left(\mathbf{x}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right)$ falling outside the domain. Thus, η_ω will be greater than one for a convex boundary and less than one for a concave boundary. Essentially, the problem of non-bijection of the mirror points is fixed through this parameter.

For Dirichlet boundary conditions:

$$\overline{W}_{\mathfrak{D}}\left(\mathbf{x}_\omega, \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right) \coloneqq \overline{W}\left(\mathbf{x}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right) + \eta_\omega \left(\frac{2\mu_{\mathrm{o}}}{\mu_\omega} - 1\right) \overline{W}\left(\widetilde{\mathbf{x}}_\omega - \mathbf{x}_u; \mathbf{h}_\omega, \boldsymbol{\lambda}\right), \tag{50}$$

with the same definition for η_ω. In the event that, because of the irregularity, the use of kernel (50) may generate negative density estimations in some bins, those values should be set to zero.

See §4.3.4 for an example implementation of the proposed approach.

4. Computational Investigations and Discussion

In this section we investigate the performance of the proposed methodology on hypothetical RWPT models, with a focus on the features that are new with respect to the original approach introduced in [17]. This includes the Eulerian grid projection, the iterative nature of the local kernel optimization, and the accounting for boundary conditions, both regular and irregular. In Appendix D we also discuss the use of local auxiliary parameters in the optimization process. We study each of these aspects separately, and then we focus on fixed-time concentration estimations rather than full reactive simulations, as it should be clear that a better concentration estimation will result in improved reactive transport modeling [17].

4.1. Grid-Projected KDE

The use of kernel functions in particle-based methodologies is common for non-linear processes that involve interaction between individual particles, such as chemical reactions. Traditionally, the representation of the support volume of a particle as a smoothing kernel results in a loop through all pairs of potentially interacting numerical particles [16]. Let us consider the simple example of an

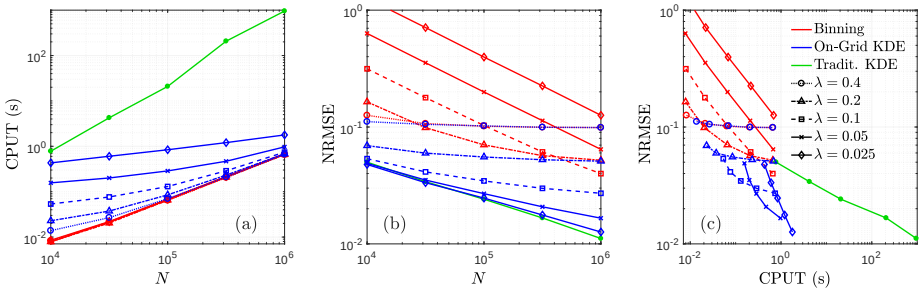

Figure 2: For a density estimation given the setup described in §4.1, using different methodologies and bin sizes: (a) CPU Time invested as a function of the number of particles, (b) Normalized Root Mean Squared Error (NRMSE) as a function of the number of particles, (c) Resulting relationship between CPUT and NRMSE.

irreversible bimolecular reaction of type,

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C}, \tag{51}$$

where A, B and C are chemical compounds, and the "batch" reaction kinetics (that is, neglecting transport) can be described as:

$$\frac{\partial c_\mathrm{C}}{\partial t} = -\frac{\partial c_\mathrm{A}}{\partial t} = -\frac{\partial c_\mathrm{B}}{\partial t} = k c_\mathrm{A} c_\mathrm{B}, \tag{52}$$

where k is the kinetic reaction constant and c_A, c_B and c_C are the concentrations of compounds A, B and C, respectively. Although chemical reactions occurring in nature - and their kinetics - are typically more complex than (51) and (52), here we use this simple setup as an illustrative example to test and compare - in terms of accuracy and performance - different RWPT approaches to simulating reactions. The probability of reaction of a particle a of compound A in a time step Δt can be estimated as [17]:

$$P = k \Delta t c_\mathrm{B}(\mathbf{X}_a) = \frac{1}{\phi} k \Delta t m_\mathrm{B} \rho_\mathrm{B}(\mathbf{X}_a). \tag{53}$$

For the sake of simplicity, ϕ in (53) is assumed constant. Here, the density of B-particles ρ_{B} is estimated from equations (1) and (2) at position $\mathbf{X}_a$ of the A-particle a. Doing this for all potentially reactive A-particles would involve a double loop to see the interactions between all particle pairs, and therefore scales in number of calculations as $N_{\mathrm{A}}N_{\mathrm{B}}$. In addition, it also requires a search algorithm, which would scale at best as $N_{\mathrm{A}} \log N_{\mathrm{B}}$. Aside from incorporating several kinds of boundary conditions, as addressed in §3, a powerful argument for performing a pilot binning before the kernel smoothing is to increase computational efficiency, by pre-grouping the particles and avoiding the need for search algorithms; additionally one can pre-compute and store the matrix kernels, thus not having to evaluate the kernel function continuously. In this way, we simultaneously benefit from the low computational demands of binning as well as the accuracy of KDE.

In order to exemplify this, consider a simple hypothetical 2D problem where, at a given time, compounds A and B, each represented by $N_{\mathrm{A}} = N_{\mathrm{B}} = N$ particles, are distributed in space as partially overlapping multi-Gaussian distributions. Both these distributions have isotropic, unit variance $\sigma^2 = 1$, and their centers $\langle \mathbf{X}_{\mathrm{A}} \rangle$, $\langle \mathbf{X}_{\mathrm{B}} \rangle$, are separated by a distance 0.8σ.

To perform a reactive time step given the described conditions, we consider three possible alternatives to estimate $\rho_{\mathrm{B}}(\mathbf{X}_a)$ in (53), for all $a = 1, \ldots, N$: (i) Binning, (ii) Traditional KDE (eq. (2)), and (iii) The KDE method proposed herein (eq. (4)).

To independently evaluate and compare these techniques, we use a constant (global), isotropic kernel bandwidth of size $\widehat{h} = \sigma N^{-\frac{1}{6}}$ (see equation (10)). In the limit of $N \to \infty$ and $\boldsymbol{\lambda} \to \mathbf{0}$, the estimated density at a point should converge

to the true solution

$$\rho_{\mathrm{B}}^{*}(\mathbf{X}_a) = \frac{N}{2\pi\sigma^2} \exp\left(-\frac{\left[\mathbf{X}_a - \langle \mathbf{X}_{\mathrm{B}} \rangle\right]^2}{2\sigma^2}\right). \tag{54}$$

We measure the difference between ρ_{B} and ρ_{B}^{*} through the normalized root mean squared error (NRMSE):

$$\mathrm{NRMSE} \coloneqq \left[\frac{\sum_{a=1}^{N} \left[\rho_{\mathrm{B}}(\mathbf{X}_a) - \rho_{\mathrm{B}}^{*}(\mathbf{X}_a)\right]^2}{\sum_{a=1}^{N} \left[\rho_{\mathrm{B}}^{*}(\mathbf{X}_a)\right]^2}\right]^{\frac{1}{2}}. \tag{55}$$

Additionally, we compute the time spent to perform the density estimation. Figure 2 compares NRMSE, N and CPU time, given different values of isotropic bin size λ.

For fixed N and λ, the proposed density estimation technique is always more accurate than binning and more efficient than the traditional KDE. As a result, given a desired degree of accuracy, or equivalently, a fixed spatial scale of interest λ, we have strong evidence that the proposed technique is the optimal choice, in terms of attainable computational efficiency, for estimating concentrations and reaction rates. For a very high value of N, as the kernel bandwidth $\widehat{h}$ becomes small in comparison to λ, the proposed method and binning converge in terms of performance and accuracy, because there is in fact no effective smoothing. Prior to reaching that point (and mostly, in areas of lower particle density), the kernels are able to successfully and efficiently make up for the lack of particles.

4.2. Local Bandwidth Optimization

Next, the local kernel bandwidth optimization and density estimation algorithm is tested in a synthetic example of advective-dispersive transport in a 2D heterogeneous porous medium (see Figure 3(a)). The spatial distribution of log hydraulic conductivity K, $Y = \log K$, in this domain of size (80×50 m)

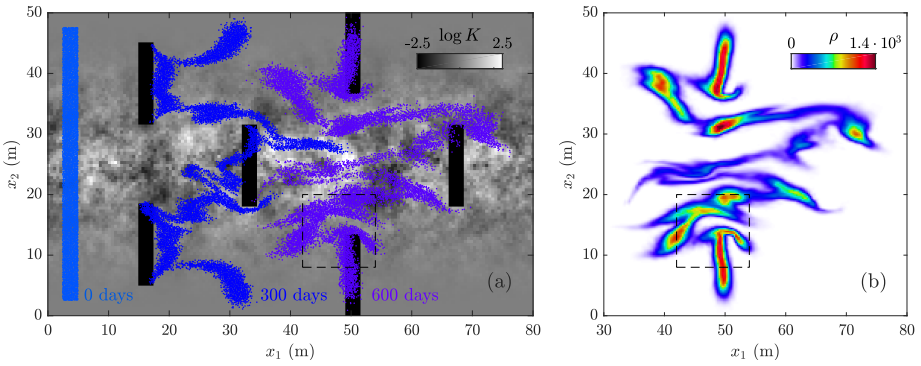

Figure 3: (a) Setup of the sample transport problem described in §4.2, and resulting particle distribution after 300 and 600 days. The dashed line signals the zoomed-in region of Figures 4 and 5. (b) Particle densities, estimated by the novel method, at $t = 600$ days.

is built in 3 steps: first, (i) a random multi-Gaussian field is generated with mean $\langle Y \rangle = 0$, variance $\langle Y^2 \rangle = 1$, and an exponential isotropic variogram with integral scale $I_Y = 3$ m; (ii) it is then multiplied with a field that evolves linearly in the vertical direction from 0 at the top and bottom boundaries to 1 at the center; (iii) finally, some low-conductivity ($Y = -2.5$) inclusions measuring 2.5×13.5 m are added as shown in Figure 3(a). The objective of this multistep procedure is to generate multiple different local features, that could be used as a test for the local bandwidth selection algorithm.

Water flows through the porous medium following:

$$\nabla \cdot \mathbf{q} = 0, \qquad \mathbf{q} = -K \nabla H, \tag{56}$$

where $\mathbf{q}$ is the Darcy velocity and H is the hydraulic head. The top and bottom boundaries are impermeable, and the head is prescribed at the left and right boundaries to force a mean hydraulic gradient of 2.5%, generating a mean flow from left to right. The domain is discretized in square cells of 0.5×0.5 m, and $\mathbf{q}$ at the cell interfaces is obtained via MODFLOW 2005 [31].

As an initial condition for transport, a uniform, rectangular distribution

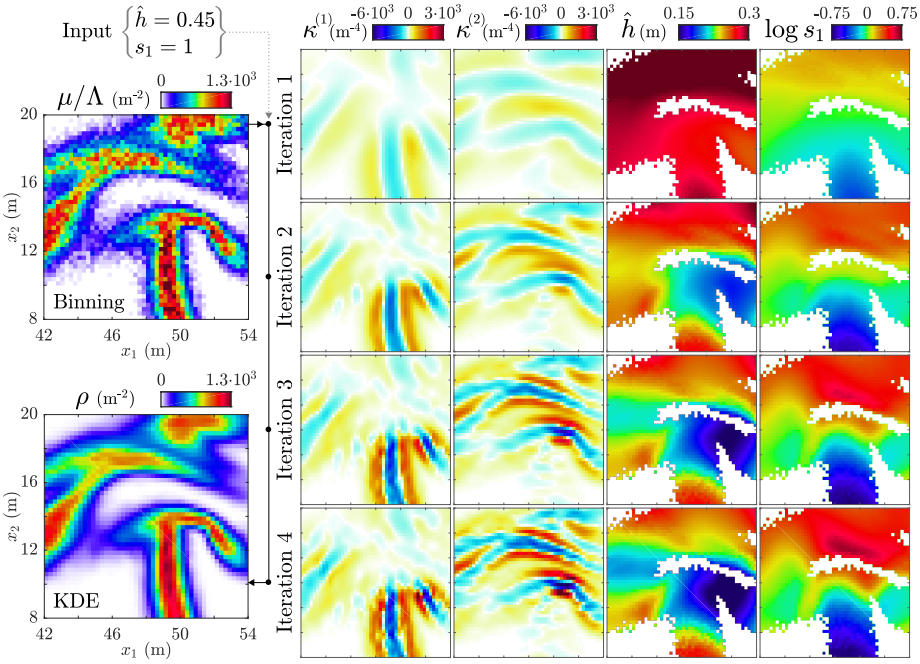

Figure 4: Graphical scheme of the kernel bandwidth optimization process, in the subregion marked by a dashed line in Figure 3, for an initial constant isotropic bandwidth of size $\widehat{h} = 0.45$. On the upper-left, the pilot binning density estimation. On the right, for each iteration, the estimated directional density curvatures, and the resulting size ($\widehat{h}$) and rightwards elongation (s_1) of the kernel bandwidth. On the lower-left, the final kernel density estimation.

of $1.8 \cdot 10^5$ particles (representing a solute) is injected at $t = 0$ near the left boundary, as shown in Figure 3. For time intervals $[t, t + \Delta t]$, particles move by random walk particle tracking (RWPT) [1]:

$$\mathbf{X}_p\left(t + \Delta t\right) = \mathbf{X}_p\left(t\right) + \mathbf{A}\left(\mathbf{X}_p\left(t\right)\right)\Delta t + \mathbf{B}\left(\mathbf{X}_p\left(t\right)\right)\boldsymbol{\xi}\sqrt{\Delta t}, \tag{57}$$

which is equivalent to solving the advection-dispersion equation (ADE). In (57), $\mathbf{A} \coloneqq \frac{1}{\phi}\left[\mathbf{q} + \nabla \cdot (\phi \mathbf{D})\right]$, with $\mathbf{D}$ being the dispersion tensor; $\mathbf{B}$ is a $d \times d$ matrix such that $\mathbf{B}\mathbf{B}^{\mathrm{T}} = 2\mathbf{D}$; and $\boldsymbol{\xi}$ is a $d \times 1$ vector of standard-normally distributed random numbers, uncorrelated in time. The spatially variable, anisotropic dispersion tensor $\mathbf{D}$ is determined following [32], with a longitudinal dispersivity of $\alpha_\ell = 2 \cdot 10^{-3}$ m, a transverse dispersivity of $\alpha_t = 3 \cdot 10^{-4}$ m, and a molecular

diffusion of $D_m = 2 \cdot 10^{-4}$ m^2/day. The spatial interpolation of velocities and the dispersion tensor is done using the hybrid linear-bilinear method proposed by [33]. Considering the heterogeneity length-scale, the Péclet number for this setting is much higher than 1 (of the order of 10^3). With this we intend to reproduce (i) an example of a scenario where one would likely choose a Lagrangian approach over a Eulerian one, and (ii) an adequate testing ground for the local kernel adaptation given the variety of elongated structures and sharp gradients displayed by the solute plume (see Figure 3). A small sensitivity analysis for the locally optimal kernel's size and shape with varying flow and transport conditions can be found in [17].

Figure 3(b) shows the spatial distribution of the particle density ρ after 600 simulated days, estimated from the particle position information using the methodology presented in §2. In Figure 4, we show the iterative process of bandwidth differentiation and convergence (see §2.5), starting from an arbitrary, uniform isotropic bandwidth with $\widehat{h} = 0.45$ m, within a zoomed-in region of the domain (delimited by the dashed lines shown in Figure 3). The estimation of the curvatures $\kappa^{(i)}$, included in Figure 4, is crucial for the correct determination of the locally optimal scale $(\widehat{h})$ and elongation (s_1) of the smoothing kernel bandwidth. We see for this particular example that, after 4 iterations, the solution has nearly stabilized. The smoothing of the "pilot" (binning) concentrations through this optimal local kernel is able to visibly reduce the noise, without generating excessive over-smoothing (see the two plots on the left of Figure 4). As a result, the NRMSE was reduced from a 93.7% to a 7.0% after the smoothing, with the NRMSE in this case being defined as in (55), but for ρ on all $\mathbf{x}_u$, and with the "true" solution in this case being approximated as the binning solution obtained for $N = 8.64 \cdot 10^7$ (~ 2.5 orders of magnitude more than the test case).

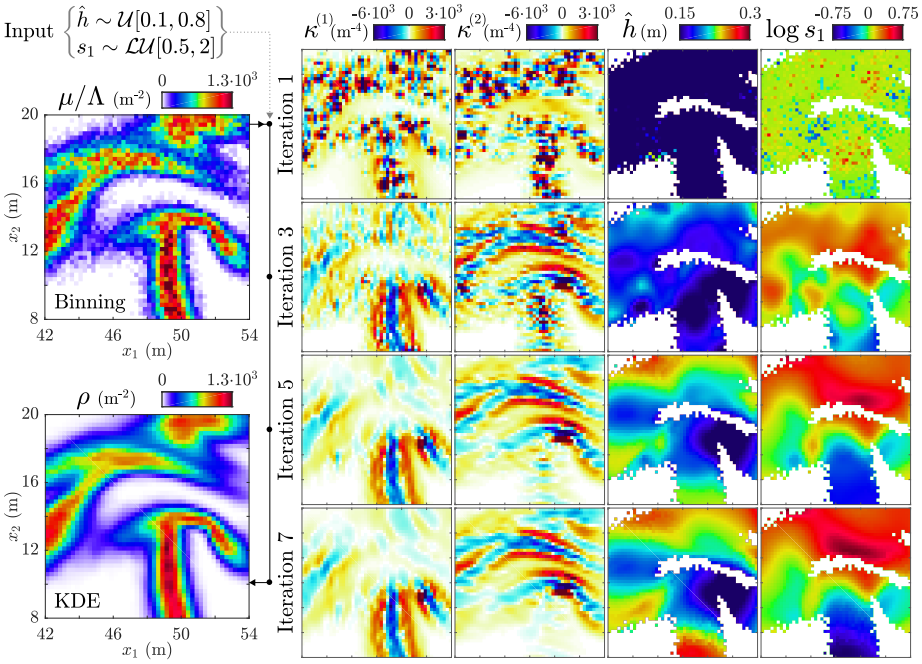

Figure 5: Graphical scheme of the kernel bandwidth optimization process, in the region marked by a dashed line in Figure 3, for an initial uncorrelated random bandwidth, with uniformly distributed size and log-uniformly distributed rightwards elongation. On the upper-left, the pilot binning density estimation. On the right, for each iteration (1,3,5,7), the estimated directional density curvatures, and the resulting size ($\widehat{h}$) and rightwards elongation (s_1) of the kernel bandwidth. On the lower-left, the final kernel density estimation.

Another relevant property of the iterative optimization algorithm is its robustness, i.e. its ability to reach the same solution given different initial values. This is particularly important in the context of particle tracking simulations, where the particles can "carry" the support volume for some number of steps and then use it as an input to update the next optimal support volume, conferring it an evolutionary nature. For this to make sense, it is critical that for any given current distribution of particles, the solution always converges to a unique set of values, regardless of the history of the local kernel bandwidth.

To test this, we repeat the previous numerical experiment, but this time, the initial bandwidth is random and uncorrelated in space, with $\widehat{h}$ drawn from a uniform distribution with lower and upper bounds of 0.1 and 0.8; and with s_1

drawn from a log-uniform distribution (the logarithm is uniformly distributed) with lower and upper limits of 0.5 and 2.0. These are deliberately chosen to be challenging for the convergence of the algorithm. As can be seen from Figure 5, in the first iteration this results in a curvature estimation $(\kappa^{(1)}, \kappa^{(2)})$ that is far from the correct solution, and characterized by high absolute values and strong variations with no apparent spatial correlation. As a consequence, the first estimation of $\widehat{h}$ and s_1 gives a small bandwidth without a well-defined elongation direction. Yet, after a few iterations, we see that the identification of the curvatures substantially improves, and concurrently the bandwidth scale and elongation start differentiating spatially distinct regions. For this highly adverse case of a locally random and uncorrelated input value, after only 7 iterations we have nearly reached the same stable solution as in Figure 4.

4.3. Boundary Conditions

Here we implement and evaluate the boundary condition correction techniques described in §3. First, we perform a concentration estimation in three simple 1D RWPT settings which are designed in such a way that there is overlap between some of the bins' uncorrected kernels and the boundary. In all cases, the solute moves by dispersion with $D = 0.1$ m^2/day for a total time of $\tau = 1000$ days; the medium porosity is $\phi = 0.25$. Each particle has a mass (or more precisely, an amount of substance) of $m = 10^{-4}$ mol. The bin size is $\lambda = 0.5$ m. We compare the three estimation methods (the pilot binning, the uncorrected KDE, and the corrected KDE) to the analytical solution for each case.

4.3.1. Dispersion near impermeable wall

In the first example, the initial condition is a Dirac delta pulse of solute (with a total mass $M = 1$ mol) located at $x = 10$ m, near an impermeable

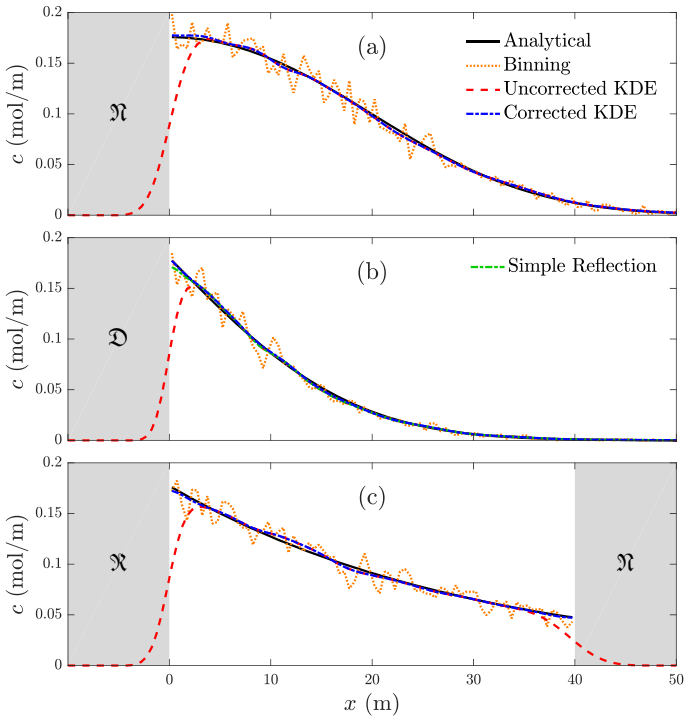

Figure 6: Comparison between the analytical and the RWPT solution, with different concentration estimation techniques, for the simple transport problems described in Section §4.3. The spurious fluctuations in the pilot binning are corrected by the optimal kernel smoothing, at the cost of inaccuracies near the boundaries. This issue is solved by the boundary corrections introduced in §3.

boundary located at $x_{\mathfrak{N}} = 0$. The results are shown in Figure 6(a). Since here diffusion is the only physical process, and the corrected kernel emulates diffusion, the concentration estimation has an excellent agreement with the analytical solution.

4.3.2. Dispersion through boundary

In the second example, all initial concentrations are zero inside the domain. There is a Dirichlet boundary condition such that $c_{\mathrm{o}} = 0.18$ mol/m at $x_{\mathfrak{D}} = 0$. The results are shown in Figure 6(b). The Dirichlet "reflection" technique (45) is able to correctly reconstruct the concentration field and gradient near the boundary. It is worth noting that a simple reflection as in (36) would have

resulted in a zero-gradient instead, as shown by the curve labeled as "Simple Reflection". Nevertheless, the error involved in using (36) instead of (45), at least in this specific case, was relatively small (compared, for instance, to the error associated with not performing a boundary correction at all).

4.3.3. Column with inlet and outlet reservoirs

In this third example we consider a constant rightwards advection ($q = 0.055$ m/day) along with a linear degradation:

$$\frac{\partial c}{\partial t} = -\frac{q}{\phi}\frac{\partial c}{\partial x} + D\frac{\partial^2 c}{\partial x^2} - kc. \tag{58}$$

The reaction is simulated stochastically by a particle reaction probability $P = k\Delta t$ (i.e., this is the probability a particle dies in any given time step). At the inlet we have a Robin boundary condition as described in §3.1.3, and at the outlet we set a zero-gradient (homogeneous Neumann) condition. This problem has a stationary solution. From Figure 6(c), we see that the corrected density estimation has a zero-gradient near the boundaries, and that the gradient is not zero in the true solution. This is because the kernel reflection method is based on pure diffusion and unable to account for the gradient generated by the combined action of advection and reaction. Nonetheless, one can appreciate in Figure 6(c) a substantial improvement for the corrected KDE method with respect to the uncorrected one.

4.3.4. Irregular Boundaries

In order to illustrate the boundary correction technique for the case of irregular boundaries explained in §3.2, we designed a problem consisting of a domain with the shape of an arched tube (Figure 7) with impermeable boundaries. The region shown has dimensions 115×115, and the spatial discretization is square with $\lambda = 1$. The particle cloud (Figure 7(a)) is the result of a random in-

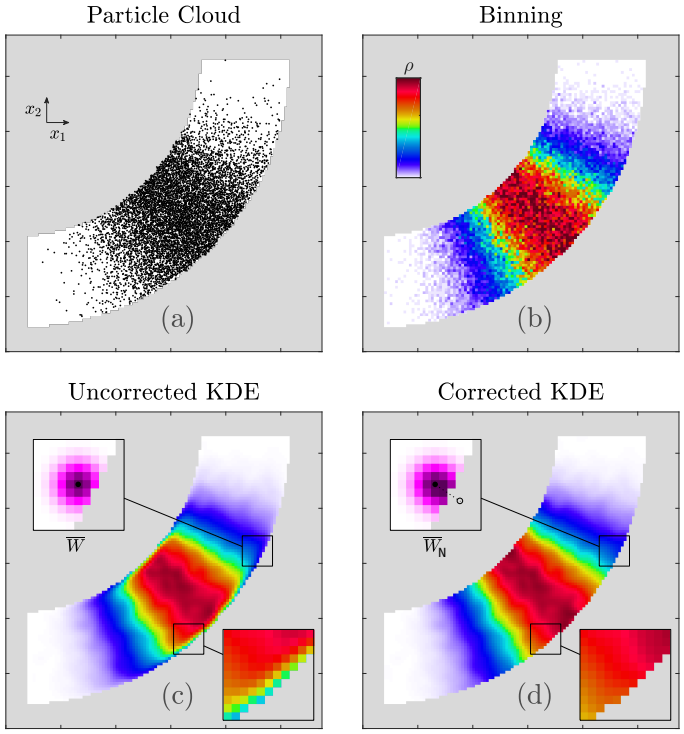

Figure 7: Graphical example of a kernel density estimation with irregular boundaries, for an impermeable boundary condition. The uncorrected KDE (c) produces an unphysical loss of mass near the boundary. This is corrected by a "mirror" modification of the kernels near the boundary (d).

jection of 100 000 particles. In polar coordinates (the origin being the center of curvature of the tube), the square of the radial coordinate of the particle positions is uniformly-distributed, whereas the angular coordinate is Gaussian-distributed. Such a distribution could be thought of as a hypothetical result of advection-dispersion through the tube.

Despite the high number of particles used, we observe in Figure 7(b) a considerable amount of noise in the binning estimation. This issue is fixed by our locally adaptive kernel smoothing technique, as shown in Figure 7(c), but at the expense (prior to any boundary correction) of an artificial density loss near the boundaries (see zoomed-in region). As illustrated by the representation of

the kernel $\overline{W}$ near the boundary, this is caused by the lack of a reflection of the mass lost through the boundary by the smoothing. In Figure 7(d), we see that the modified kernel (see zoomed-in representation) successfully fixes the loss of mass through the boundary, yielding a satisfactory reconstruction of the concentrations.

5. Summary and Conclusions

We have presented a novel technique to estimate particle densities using the limited amount of information provided by a finite sample of particle positions. Although the spectrum of possible applications is wide, the focus of this work is the reconstruction of solute concentrations in random walk particle tracking (RWPT) simulations, which is relevant for the visualization of results and the incorporation of chemical reactions. Our technique relies on the accuracy of locally adaptive kernel density estimation (KDE), which is implemented in combination with a spatial discretization, resulting in benefits including computational efficiency and accurate implementation of boundary conditions. The method is valid in 1, 2 or 3 spatial dimensions. In principle, it can be applied to Lagrangian modeling techniques other than RWPT, or even to other applications of density estimation. An open-source MATLAB code, "Bounded Adaptive Kernel Smoothing" (BAKS), has been developed and made available to the community as a result of this work [34].

Our computational investigations provide strong evidence that our proposed methodology deals well with the dilemma between the accuracy of kernel methods and the low computational requirements of binning. From our tests, we see that the desired degree of accuracy (characterized by the choice of bin size $\boldsymbol{\lambda}$) is always achieved faster (in terms of invested computational effort) with our proposed methodology compared with binning. The two methods converge, both in

accuracy and performance, for high particle numbers, but prior to reaching that point, the kernel approach reaches an acceptable level of error earlier. Likewise, direct Lagrangian utilization of the kernel functions involves a worse scaling of CPU time with particle number than our method. As a result, the proposed method can achieve the same level of accuracy with lower computational effort.

Through a sample simulation of conservative transport in a heterogeneous porous medium, we show the convergence of the local kernel bandwidth iterative optimization method towards a stable result. Using that optimal local bandwidth to smooth the pilot binning density estimation resulted in a reduction of normalized error from 93.7% to 7.0%. We also demonstrate the robustness of the method, in terms of being able to reach this same solution regardless of the initial input values provided. The fully local nature of the optimization process, which is a novel aspect with respect to the original methodology [17], allows the kernel to achieve a higher degree of local differentiation in terms of size and shape.

Near boundaries, the kernel is assumed to emulate the Green's function of a fast diffusion process. Hence, the kernel is affected by the boundary following the analytical solution of pure diffusion corresponding to the relevant boundary condition, which results in simple and efficient reflection rules. The case of irregular boundaries is slightly more complicated, as it requires individual modification of each kernel, with a weighting of the reflection to ensure proper mass conservation. The simple implementation examples illustrate the importance of including these boundary corrections in the density estimation.

Appendix A. Details on treatment of matrix kernels

The projected kernels $\overline{W}$ and $\overline{V}$, introduced in §2.1 and §2.2, respectively, require some corrections to ensure that they keep the main properties of the

original kernels W and V after the on-grid projection. In the case of $\overline{W}$, the original Gaussian kernel W integrates exactly to 1 only over $\mathbb{R}^d$. If a cutoff distance is imposed, a normalization is needed to impose mass conservation:

$$\overline{W}'\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{h},\boldsymbol{\lambda}\right) \coloneqq \frac{\overline{W}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{h},\boldsymbol{\lambda}\right)}{\sum_{\boldsymbol{\zeta}}\overline{W}\left(\boldsymbol{\lambda}\odot\boldsymbol{\zeta};\mathbf{h},\boldsymbol{\lambda}\right)}, \tag{A.1}$$

where $\overline{W}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{h},\boldsymbol{\lambda}\right)$ is the unmodified kernel at cell index $\mathbf{z}$ (equation (8)), $\mathbf{z}=\mathbf{0}$ corresponding to the cell where the center of the kernel is located, and $\overline{W}'$ is the modified kernel. In (A.1), the denominator is a sum over all cells within the cutoff limits.

In the case of $\overline{V}$, there are two corrections that need to be performed in order to conserve the original purpose of V. On one hand, the positive values must be weighted so that the kernel integrates to zero within the cutoff limits:

$$\overline{V}'^{(i)}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{g},\boldsymbol{\lambda}\right) \coloneqq a\overline{V}^{(i)}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{g},\boldsymbol{\lambda}\right), \tag{A.2}$$

$$a \coloneqq \begin{cases} -\dfrac{\sum_{\overline{V}^{(i)}<0}\overline{V}^{(i)}\left(\boldsymbol{\lambda}\odot\boldsymbol{\zeta};\mathbf{g},\boldsymbol{\lambda}\right)}{\sum_{\overline{V}^{(i)}>0}\overline{V}^{(i)}\left(\boldsymbol{\lambda}\odot\boldsymbol{\zeta};\mathbf{g},\boldsymbol{\lambda}\right)}, & \text{if } \overline{V}^{(i)}>0 \\ 1, & \text{if } \overline{V}^{(i)}\leq 0 \end{cases}. \tag{A.3}$$

A simple intuitive example of the importance of this correction is that a constant particle density must necessarily yield a zero-curvature estimation, and this will only occur if the curvature kernel has zero-mean.

Besides, the final purpose of $\overline{V}^{(i)}$ is the estimation of the squared second spatial derivatives of the particle density ($\kappa_\omega^{(i)}\kappa_\omega^{(j)}$ in (14)). The spatial averaging involved in the grid projection could lead to a systematic under-prediction of these values. To prevent that, we scale the kernel so that it keeps the L^2-norm

of the original kernel V after the projection:

$$\overline{V}''^{(i)}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{g},\boldsymbol{\lambda}\right)\coloneqq\left[\frac{\Lambda\left\|V^{(i)}\right\|^{2}}{\sum_{\boldsymbol{\zeta}}\left[\overline{V}'^{(i)}\left(\boldsymbol{\lambda}\odot\boldsymbol{\zeta};\mathbf{g},\boldsymbol{\lambda}\right)\right]^{2}}\right]^{\frac{1}{2}}\cdot\overline{V}'^{(i)}\left(\boldsymbol{\lambda}\odot\mathbf{z};\mathbf{g},\boldsymbol{\lambda}\right),\quad\text{(A.4)}$$

where $\|\ \|^2$ is the L^2-norm operator, i.e.,

$$\left\|V^{(i)}\right\|^{2}\coloneqq\int_{\mathbb{R}^{d}}\left[V^{(i)}\left(\mathbf{r};\mathbf{g}\right)\right]^{2}\mathrm{d}\mathbf{r}=\frac{3}{2^{(d+2)}\pi^{\frac{d}{2}}g_{i}^{5}\left(\prod_{j\neq i}g_{j}\right)}.\quad\text{(A.5)}$$

$\overline{V}''^{(i)}$ are the final values of the projected curvature kernel.

As mentioned in the main body of this work, the values of $\overline{W}$ and $\lambda_i^2\overline{V}^{(i)}$ only depend on the directional ratios between the bandwidth ($\mathbf{h}$ or $\mathbf{g}$) and the grid size ($\boldsymbol{\lambda}$). Discretizing the values that these ratios are allowed to adopt, and imposing a cutoff distance (as in Figure 1), $\overline{W}$ and $\lambda_i^2\overline{V}^{(i)}$ become finite sets of matrices with a finite number of entries. Repeated use of the same (or very similar) kernel bandwidth results in the exact same matrix kernel, which can be stored in the memory after its first generation and correction. Then, computation of (4), (17), (12) and (14) only requires accessing the pre-computed matrix entries in the memory and performing the relevant weighted summation, avoiding redundant computational efforts.

Appendix B. Derivation of expression (22)

Within a virtual Gaussian distribution of N_u^σ particles, centered at $\boldsymbol{\mu}_u\equiv[\mu_{u,1},\ldots,\mu_{u,d}]^{\mathrm{T}}$ with the vector of directional standard deviations $\boldsymbol{\sigma}_u\equiv[\sigma_{u,1},\ldots,\sigma_{u,d}]^{\mathrm{T}}$,

$$\rho(\mathbf{x})=N_u^\sigma\prod_{i=1}^{d}\frac{1}{\sqrt{2\pi}\sigma_{u,i}}\exp\left(-\frac{(x_i-\mu_{u,i})^2}{2\sigma_{u,i}^2}\right)\equiv N_u^\sigma W\left(\mathbf{x}-\boldsymbol{\mu}_u;\boldsymbol{\sigma}_u\right).\quad\text{(B.1)}$$

If $\boldsymbol{\sigma}_u$ is also the bandwidth of the integration kernel, then following (12),

$$\begin{aligned} n_u &= \int \rho\left(\mathbf{x}\right) W\left(\mathbf{x}-\mathbf{x}_u;\boldsymbol{\sigma}_u\right) \mathrm{d}\mathbf{x} \\ &= N_u^\sigma \int W\left(\mathbf{x}-\boldsymbol{\mu}_u;\boldsymbol{\sigma}_u\right) W\left(\mathbf{x}-\mathbf{x}_u;\boldsymbol{\sigma}_u\right) \mathrm{d}\mathbf{x} \\ &= N_u^\sigma W\left(\mathbf{x}_u-\boldsymbol{\mu}_u;\sqrt{2}\boldsymbol{\sigma}_u\right) = \frac{N_u^\sigma [W\left(\mathbf{x}_u-\boldsymbol{\mu}_u;\boldsymbol{\sigma}_u\right)]^{1/2}}{(8\pi\widehat{\sigma}_u^2)^{d/4}} \\ &= \sqrt{\frac{N_u^\sigma \rho(\mathbf{x}_u)}{(\sqrt{8\pi}\widehat{\sigma}_u)^d}} \approx \sqrt{\frac{N_u^\sigma \rho_u}{(\sqrt{8\pi}\widehat{\sigma}_u)^d}}. \end{aligned} \tag{B.2}$$

Here, the approximation $\rho(\mathbf{x}_u) \approx \rho_u$ is equivalent to the one in (12). Taking the square on both sides of expression (B.2) and rearranging we obtain

$$N_u^\sigma = \frac{\left(\sqrt{8\pi}\widehat{\sigma}_u\right)^d n_u^2}{\rho_u}, \tag{B.3}$$

which is the expression given in (22).

Appendix C. Curvature kernel bandwidth for Gaussian distribution

Following the approach of [25, Appendix E], here generalized for d dimensions, the optimal isotropic $\widehat{g}^{(i)}$, given a distribution of particles $\rho(\mathbf{x})$, will be

$$\widehat{g}^{(i)} = \left(\frac{(2+2^{-\frac{d}{2}-1})N}{(2\pi)^{\frac{d}{2}} \sum_{j=1}^d \int \left(\frac{\partial^3 \rho}{\partial x_i^2 \partial x_j}\right)^2 \mathrm{d}\mathbf{x}} \right)^{\frac{1}{d+6}}, \tag{C.1}$$

with N being the total number of particles. Taking $\rho(\mathbf{x})$ to be a Gaussian distribution of N particles with a vector of directional standard deviations such that $\boldsymbol{\sigma} = \widehat{\sigma}\mathbf{s}$, with $\prod_{i=1}^d s_i = 1$, we have

$$\int \left(\frac{\partial^3 \rho}{\partial x_i^2 \partial x_j}\right)^2 \mathrm{d}\mathbf{x} = \frac{3(1+4\delta_{ij})N^2}{8(4\pi)^{\frac{d}{2}} \widehat{\sigma}^{d+6} s_i^4 s_j^2}, \tag{C.2}$$

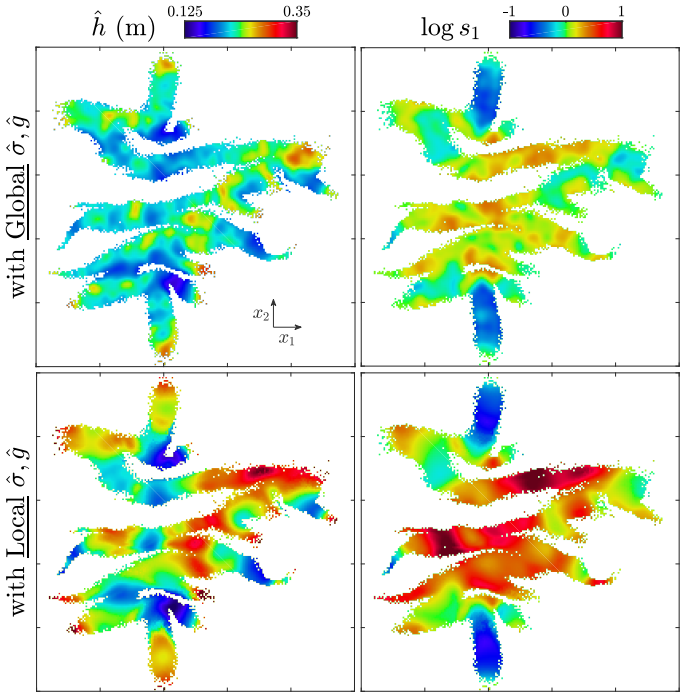

Figure D.8: Graphical comparison of the size $\widehat{h}$ (first column) and rightwards elongation s_1 (second column) of the kernel bandwidth obtained by optimization using a global integration kernel $\widehat{\sigma}$ and curvature kernel $\widehat{g}$ (first row), or local values instead (second row). Note the higher degree of local differentiation in the second case.

and then substituting (C.2) in (C.1) and rearranging we obtain

$$\widehat{g}^{(i)} = \left[\frac{4 + 2^{\frac{d}{2}+4}}{3N \sum_{j=1}^{d} \frac{1+4\delta_{ij}}{s_i^4 s_j^2}} \right]^{\frac{1}{d+6}} \widehat{\sigma}, \tag{C.3}$$

which is equivalent to the combination of equations (25) and (26) given in §2.4.

Appendix D. Local vs Global selection of σ and g

Within the procedure presented in §2, the selection of the curvature kernel bandwidth ($\mathbf{g}_u^{(i)}$) at a bin u is completely local and independent of all particles located outside the range determined by the integration support $\boldsymbol{\sigma}_u$. In the original methodology presented in [17], a global $\mathbf{g}^{(i)}$ was used over the whole

domain, as noted in §2.4. As a result, the local kernel bandwidth would be indirectly conditioned by global features. For instance, if the plume was dominantly formed by strongly elongated shapes in one specific direction, the curvature estimation anywhere would be biased towards detecting those kinds of features. This would limit the ability of the kernel bandwidth to optimally adapt strictly to the nearby particle distribution. To illustrate this, we compare the distribution of local bandwidths $\mathbf{h}$ obtained with the method presented in §2 to the one that we obtain when using global (instead of local) values for $\boldsymbol{\sigma}$ and $\mathbf{g}$, following what is described in [17]. We use the same example particle distribution of §4.2.

Figure D.8 shows the local bandwidth scale ($\widehat{h}$) and elongation (s_1) values obtained with the two described approaches. We clearly observe a higher degree of differentiation, that is, an increased ability of the bandwidth scale and elongation to reach extreme values when using the novel, fully local methodology presented here. This is particularly true for large bandwidths, which can only exist in the absence of noise in the curvature estimation (see Figure 5 as an example). The distribution of bandwidth scales and elongations also appears to be significantly smoother (less affected by "spurious" fluctuations) with the novel methodology, which suggests a more accurate identification of the dominant shapes of the local particle distribution.

Quantitatively, the NRMSE of ρ changes from 7.4% for the global parameter choice to the aforementioned 7.0% when using the novel fully local methodology. Although this may seem like a small reduction, the difference would probably become larger in the case of an even stronger spatial differentiation of the local particle distributions.

References

[1] P. Salamon, D. Fernàndez-Garcia, J. J. Gómez-Hernández, A review and numerical assessment of the random walk particle tracking method, Journal of Contaminant Hydrology 87 (3-4) (2006) 277–305. doi:10.1016/j.jconhyd.2006.05.005.

[2] P. Salamon, D. Fernàndez-Garcia, J. J. Gómez-Hernández, Modeling mass transfer processes using random walk particle tracking, Water Resources Research 42 (11). doi:10.1029/2006WR004927.

[3] B. Berkowitz, A. Cortis, M. Dentz, H. Scher, Modeling Non-fickian transport in geological formations as a continuous time random walk, Reviews of Geophysics 44 (2). doi:10.1029/2005RG000178.

[4] D. A. Benson, T. Aquino, D. Bolster, N. Engdahl, C. V. Henri, D. Fernàndez-Garcia, A comparison of Eulerian and Lagrangian transport and non-linear reaction algorithms, Advances in Water Resources 99 (2017) 15–37. doi:10.1016/j.advwatres.2016.11.003.

[5] N. B. Engdahl, M. J. Schmidt, D. A. Benson, Accelerating and parallelizing lagrangian simulations of mixing-limited reactive transport, Water Resources Research 55. doi:10.1029/2018WR024361.

[6] D. A. Benson, M. M. Meerschaert, Simulation of chemical reaction via particle tracking: Diffusion-limited versus thermodynamic rate-limited regimes, Water Resources Research 44 (12). doi:10.1029/2008WR007111.

[7] A. Paster, D. Bolster, D. A. Benson, Particle tracking and the diffusion-reaction equation, Water Resources Research 49 (1) (2013) 1–6. doi:10.1029/2012WR012444.

[8] D. Bolster, A. Paster, D. A. Benson, A particle number conserving Lagrangian method for mixing-driven reactive transport, Water Resources Research 52 (2) (2016) 1518–1527. doi:10.1002/2015WR018310.

[9] L. J. Perez, J. J. Hidalgo, M. Dentz, Upscaling of mixing-limited bimolecular chemical reactions in poiseuille flow, Water Resources Research 55 (1) (2019) 249–269. doi:10.1029/2018WR022730.

[10] A. Paster, D. Bolster, D. A. Benson, Connecting the dots: Semi-analytical and random walk numerical solutions of the diffusion-reaction equation with stochastic initial conditions, Journal of Computational Physics 263 (2014) 91–112. doi:10.1016/j.jcp.2014.01.020.

[11] D. Ding, D. A. Benson, D. Fernàndez-Garcia, C. V. Henri, D. W. Hyndman, M. S. Phanikumar, D. Bolster, Elimination of the Reaction Rate "Scale Effect": Application of the Lagrangian Reactive Particle-Tracking Method to Simulate Mixing-Limited, Field-Scale Biodegradation at the Schoolcraft (MI, USA) Site, Water Resources Research 53 (12) (2017) 10411–10432. doi:10.1002/2017WR021103.

[12] M. Rahbaralam, D. Fernàndez-Garcia, X. Sanchez-Vila, Do we really need a large number of particles to simulate bimolecular reactive transport with random walk methods? A kernel density estimation approach, Journal of Computational Physics 303 (2015) 95–104. doi:10.1016/j.jcp.2015.09.030.

[13] D. A. Benson, D. Bolster, Arbitrarily complex chemical reactions on particles, Water Resources Research 52 (11) (2016) 9190–9200. doi:10.1002/2016WR019368.

[14] N. B. Engdahl, D. A. Benson, D. Bolster, Lagrangian simulation of mixing and reactions in complex geochemical systems, Water Resources Research 53 (4) (2017) 3513–3522. doi:10.1002/2017WR020362.

[15] P. A. Herrera, M. Massabó, R. D. Beckie, A meshless method to simulate solute transport in heterogeneous porous media, Advances in Water Resources 32 (3) (2009) 413–429. doi:10.1016/j.advwatres.2008.12.005. URL http://linkinghub.elsevier.com/retrieve/pii/S0309170808002273

[16] G. Sole-Mari, D. Fernàndez-Garcia, P. Rodríguez-Escales, X. Sanchez-Vila, A KDE-Based Random Walk Method for Modeling Reactive Transport With Complex Kinetics in Porous Media, Water Resources Research- doi:10.1002/2017WR021064.

[17] G. Sole-Mari, D. Fernàndez-Garcia, Lagrangian modeling of reactive transport in heterogeneous porous media with an automatic locally adaptive particle support volume, Water Resources Research 54 (10) (2018) 8309–8331. doi:10.1029/2018WR023033.

[18] P. Szymczak, A. J. C. Ladd, Boundary conditions for stochastic solutions of the convection-diffusion equation, Phys. Rev. E 68 (2003) 036704. doi:10.1103/PhysRevE.68.036704.

[19] P. Szymczak, A. J. C. Ladd, Stochastic boundary conditions to the convection-diffusion equation including chemical reactions at solid surfaces, Phys. Rev. E 69 (2004) 036704. doi:10.1103/PhysRevE.69.036704.

[20] J. Koch, W. Nowak, A method for implementing dirichlet and third-type boundary conditions in ptrw simulations, Water Resources Research 50 (2) 1374–1395. arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013WR013796, doi:10.1002/2013WR013796.

[21] G. Boccardo, I. M. Sokolov, A. Paster, An improved scheme for a robin boundary condition in discrete-time random walk algo-

rithms, Journal of Computational Physics 374 (2018) 1152 – 1165. doi:https://doi.org/10.1016/j.jcp.2018.08.009.

[22] D. Fernàndez-Garcia, X. Sanchez-Vila, Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions, Journal of Contaminant Hydrology 120-121 (C) (2011) 99–114. doi:10.1016/j.jconhyd.2010.05.001.

[23] D. Pedretti, D. Fernàndez-Garcia, An automatic locally-adaptive method to estimate heavily-tailed breakthrough curves from particle distributions, Advances in Water Resources 59 (2013) 52–65. doi:10.1016/j.advwatres.2013.05.006.

[24] B. W. Silverman, Density Estimation for Statistics and Data Analysis, Vol. 37, 1986. doi:10.2307/2347507.

[25] Z. I. Botev, J. F. Grotowski, D. P. Kroese, Kernel density estimation via diffusion, Ann. Statist. 38 (5) (2010) 2916–2957. doi:10.1214/10-AOS799.

[26] J. S. Marron, D. Ruppert, Transformations to reduce boundary bias in kernel density estimation, Journal of the Royal Statistical Society. Series B (Methodological) 56 (4) (1994) 653–671.

[27] G. Sole-Mari, M. J. Schmidt, S. D. Pankavich, D. A. Benson, Numerical equivalence between sph and probabilistic mass transfer methods for lagrangian simulation of dispersion, Advances in Water Resources 126 (2019) 108 – 115. doi:https://doi.org/10.1016/j.advwatres.2019.02.009.

[28] S. Chandrasekhar, Stochastic problems in physics and astronomy, Rev. Mod. Phys. 15 (1943) 1–89. doi:10.1103/RevModPhys.15.1.

[29] M. T. van Genuchten, W. J. Alves, Analytical solutions of the one-dimensional convective-dispersive solute transport equation, Technical Bul-

letins 157268, United States Department of Agriculture, Economic Research Service (1982).

[30] G. Boccardo, I. M. Sokolov, A. Paster, An improved scheme for a robin boundary condition in discrete-time random walk algorithms, Journal of Computational Physics 374 (2018) 1152 – 1165. doi:https://doi.org/10.1016/j.jcp.2018.08.009.

[31] A. W. Harbaugh, MODFLOW-2005 , The U . S . Geological Survey Modular Ground-Water Model — the Ground-Water Flow Process, U.S. Geological Survey Techniques and Methods (2005) 253doi:U.S. Geological Survey Techniques and Methods 6-A16.

[32] J. Bear, A. Cheng, Modeling Groudwater flow and contaminant transport, in: Theory and Applications of Transport in Porous Media, 2010, p. 850. arXiv:9809069v1, doi:10.1007/978-1-4020-6682-5.

[33] E. M. LaBolle, G. E. Fogg, A. F. B. Tompson, Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods, Water Resources Research 32 (3) (1996) 583–593. doi:10.1029/95WR03528.

[34] G. Sole-Mari, Bounded Adaptive Kernel Smoothing (BAKS) [Software] doi:10.5281/zenodo.2762790.